

100-205711
20521
P-19

Experimentation in Machine Discovery

DEEPAK KULKARNI

AI Research Branch, Mail Stop 244-17
NASA Ames Research Center
Moffett Field, CA 94035

HERBERT A. SIMON

Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213

(NASA-IM-107874) EXPERIMENTATION IN MACHINE
DISCOVERY (NASA) 19 p

N92-25740

Unclas
G3/63 0091521

NASA Ames Research Center
Artificial Intelligence Research Branch

Technical Report FIA-90-09-24-02

September 1, 1990

Supersede FIA-90-09-24-01

REPORT DOCUMENTATION PAGE

Form Approved

OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE Dates attached	3. REPORT TYPE AND DATES COVERED	
4. TITLE AND SUBTITLE Titles/Authors - Attached			5. FUNDING NUMBERS	
6. AUTHOR(S)				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Code FIA - Artificial Intelligence Research Branch Information Sciences Division			8. PERFORMING ORGANIZATION REPORT NUMBER Attached	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Nasa/Ames Research Center Moffett Field, CA. 94035-1000			10. SPONSORING/MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION /AVAILABILITY STATEMENT Available for Public Distribution <i>Pete Fiedler</i> 5/14/92 BRANCH CHIEF			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) Abstracts ATTACHED				
14. SUBJECT TERMS			15. NUMBER OF PAGES	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT	

Experimentation in Machine Discovery

DEEPAK KULKARNI

AI Research Branch, Mail Stop 244-17
NASA Ames Research Center
Moffett Field, CA 94035

HERBERT A. SIMON

Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213

Abstract

This report describes KEKADA, a system that is capable of carrying out a complex series of experiments on problems from the history of science. The system incorporates a set of experimentation strategies that were extracted from the traces of scientists' behaviors. It focuses on surprises to constrain its search, and uses its strategies to generate hypotheses and to carry out experiments. Some strategies are domain independent, whereas others incorporate knowledge of a specific domain. The domain-independent strategies include magnification, determining scope, divide and conquer, factor analysis, and relating different anomalous phenomena. KEKADA represents an experiment as a set of independent and dependent entities, with apparatus variables and a goal. It represents a theory either as a sequence of processes or as abstract hypotheses. This report describes KEKADA's response to a particular problem in biochemistry. On this and other problems, the system is capable of carrying out a complex series of experiments to refine domain theories. Analysis of the system and its behavior on a number of different problems has established its generality, but it has also revealed the reasons why the system, in its present form, would not be a good experimental scientist.

This report appeared as a chapter in J. Shrager & P. Langley (Eds.) (1990), *Computational models of scientific discovery and theory formation*. San Mateo, CA: Morgan Kaufmann.

1. Introduction

Although experimentation plays an important role in many fields of science, only a few studies have focused on the role of experimentation in scientific discoveries (Friedland, 1979; Karp, 1989; Rajamoney, 1990). This chapter describes KEKADA, a system capable of carrying out a complex series of experiments.

The study of surprising phenomena is an important research task in many experimental domains. This task can be formulated as:

- *Given:* A surprising phenomenon;
- *Given:* A set of facts and hypotheses about the domain;
- *Do:* Carry out an experimentation program, acquire the results of the experiments, and use these results to refine the domain theory incrementally.

Unlike the programs developed by Karp (1990) and Rajamoney (1990), KEKADA is capable of detecting surprising phenomena and using these surprises to guide its attempts to revise the domain theory. In the next section, we describe the representation used in the system. After this, we describe the program's control and its basic processes. In the fourth section, we present an example of KEKADA's behavior on a particular problem. Then in the fifth section, we analyze the abilities of the system, after which we draw some tentative conclusions.

2. Representation in KEKADA

We will first describe KEKADA's representation of its knowledge, including the experiments it suggests and the theories it refines. An experiment is an operation carried out under well-specified conditions to determine an unknown effect. For example, one may immerse liver tissue slices in a solution of ornithine, maintained at pH 8, and carry out certain tests to measure the results. A substance like ornithine, which the experimenter includes in the experimental setup, is called an *independent entity*. An experiment may also result in production of new *dependent entities*.

Each entity has an associated set of independent, dependent, or apparatus variables. For example, ornithine might have *concentration* as an independent variable and *rate of consumption* as a dependent variable. An apparatus variable plays an auxiliary role in an experiment and is not a direct cause of the results of the experiment. In the scenario that we described earlier, the experimenter uses the tissue slice method and maintains the pH at 8. In this case, both *pH* and *method* are apparatus variables. Table 1 shows KEKADA's representation of this particular experiment.

Table 1. Representation of an experiment in KEKADA.

Current goal:	Determine relevant factors of a phenomenon.
Independent entity:	Name = Ornithine, Concentration = Medium.
Apparatus variables:	Method = Tissue-slice, pH = 8.
Measure:	Rates of production of the outputs of the reaction.
Dependent entity:	Expected-Name = Urea, Expected-Rate = < 2, 10 >.

In addition to the different entities and variables, this experiment has as a goal to find relevant factors in a previously observed phenomenon. Before any experiment is carried out, KEKADA associates expectations about the values of one or more dependent entities and associated variables. For numeric attributes, lower and upper bounds on the value are specified. For symbolic attributes, a single nominal value is specified. In the experiment in Table 1, the system expects urea to be produced at a rate between 2 and 10.

Whereas KEKADA represents experiments in such specific terms, it represents hypotheses at various levels of abstraction. A hypothesis is described as a sequence of processes, where each process has the same basic representation as an experiment. For example, consider a process that is a sequence of three reactions: P1, P2, and P3. Suppose that in P1, the substances ammonia, carbon dioxide, and ornithine combine to produce citrulline and water. In P2, citrulline and ammonia combine to produce arginine and water. In P3, arginine and water combine to produce ornithine and urea. The representation of this hypothesis is shown in Table 2.

Hypotheses can be even more abstract than the one shown in Table 2. Thus, one hypothesis might specify that ornithine donates the amino group to urea in a reaction, and another hypothesis might specify that ornithine acts as a catalyst, as in:

Type: donates-group, Donor: Ornithine, Group: Amino, Receiver: Urea.

Type: is-catalyst, Reactant: Ornithine.

Each hypothesis in KEKADA has an associated confidence vector, which is represented as a 5-tuple <success, failure, failed-effort, implied-success, implied-failure>. The success slot stores the number of experiments that have verified a hypothesis, whereas the failure slot stores the number of experiments that have failed to support a hypothesis. The implied-success slot stores the number of experiments that are a positive but inconclusive indication of the validity of a hypothesis. The failed-effort slot stores the amount of effort spent to find positive instances of an existential hypothesis. Finally, the implied-failure slot stores the number of experiments that indicate, but not conclusively, that a hypothesis is false.

Table 2. Representation of an hypothesis in KEKADA.

Process: Ornithine Cycle consists of P1, P2, and P3.
Process Name: P1
Independent Entities: Ammonia, Carbon dioxide, Ornithine
Dependent Entities: Citrulline, Water
Process Name: P2
Independent Entities: Citrulline, Ammonia
Dependent Entities: Arginine, Water
Process Name: P3
Independent Entities: Arginine, Water
Dependent Entities: Ornithine, Urea

3. Control Structure and Processes

In this section, we will examine KEKADA's control structure and basic processes. The control structure incorporates two high-level techniques: *heuristic search through two problem spaces and use of surprises to direct the search*. We first discuss these techniques and then describe various components in the system.

3.1 Dual Space Search and Surprises

The basic source of KEKADA's new knowledge is the environment. The system carries out experiments on the external world to gather new information and to modify confidences in existing beliefs. Thus it searches two spaces, one containing hypotheses and the other containing experiments and results (Simon & Lea, 1974). On the basis of the current state of the hypothesis space (existing hypotheses and their confidences), the system chooses an experiment to carry out. It then interprets the outcome of the experiment, modifying its hypotheses and their confidences. Figure 1 shows this organization in a graphic form.

Let us consider an example of how KEKADA uses experiments. If the system is studying the hypothesis that a specific substance, ornithine, is acting as a catalyst in a given reaction, it may decide to carry out an experiment to verify this belief. Thus it may measure the amount of urea produced in the presence of a small amount of ornithine. If a large amount of urea is produced in this experiment, this lends evidence to the hypothesis.

KEKADA focuses its attention on surprises to constrain its search. Surprises have played a central role in many important discoveries. For example, in the course of years of research that produced

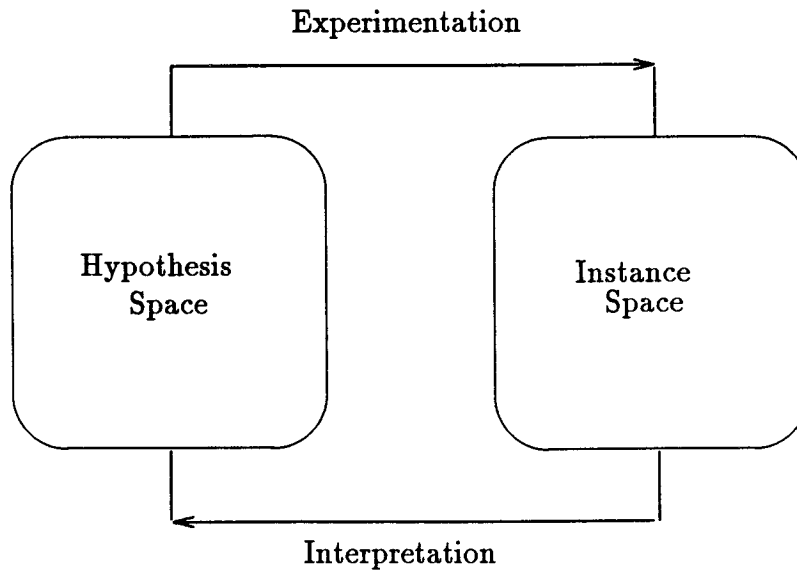


Figure 1. KEKADA's dual space search organization.

many important results, Priestley observed that “the first hints, at least of almost everything I have discovered, of much importance, have occurred to me in this manner (as unexpected phenomena)” (Conant, 1957). KEKADA attends to surprises, thereby searching the parts of its experiment space dense with useful phenomena. To this end, it associates expectations with each experiment. Thus, if it is carrying out an experiment on ornithine and ammonia in the liver tissue slices, its prior experiences may lead it to expect that the experiment would produce urea at some rate between 2 and 10 units. If urea is produced at the rate of 20 units, the outcome violates expectations. In this case, KEKADA lacked a whole body of knowledge and thus had wrong expectations about the results of the experiment. Focusing attention on this surprise puts a powerful constraint on the system's search for new knowledge.

3.2 Overall Control Structure

The overall control structure of KEKADA,¹ shown in Table 3, allows it to carry out this dual space search, to detect surprises, and to focus on them. The system has a set of strategies, each with three associated parts: *hypothesis generators*, *experiment proposers*, and *experiment evaluators*. Each of these consists of a set of rules in the form of conditions and actions. The input to KEKADA is a surprising phenomenon: the expectations and the results of an experiment. First, the hypothesis

1. KEKADA is implemented in the production system language OPS5 (Brownston, Farrell, Kant, & Martin, 1985).

A production system consists of two main components: a set of condition-action rules, or *productions*, and a dynamic working memory.

Table 3. The control structure of KEKADA.

Inputs:
R: the results of the experiment
EXP: the expectations for the experiment
Procedure KEKADA (R, EXP)
Generate a set of hypotheses HL using the hypothesis generators whose conditions are satisfied.
Repeat
Choose hypothesis HS from the set HL.
Let S be the strategy whose hypothesis generator suggested the hypothesis HS.
Generate a set of experiments EL using the experiment proposers associated with the strategy S.
For each experiment E in the set EL,
Generate expectations EXP' for the experiment E.
Carry out the experiment E to get results (R').
If results R' are not within the expectation EXP',
Then KEKADA (R', EXP');
Else interpret the results of the experiment E using the experiment evaluators associated with the hypothesis HS.

generators whose conditions are satisfied produce hypotheses. Then the system chooses one of these hypotheses using a preference scheme described in Kulkarni (1988). The strategy whose hypothesis generator suggested this hypothesis also has components for proposing experiments and evaluating them. Next the experiment proposers suggest a number of experiments, and a set of *expectation setters* generates expectations for them. At this stage, the user provides the system with the results of these experiments. If these results violate the expectations for the experiment, KEKADA detects this as a surprise and makes a recursive call to itself. If no surprise is detected, then the experiment evaluators interpret the results of the experiment. The system repeats this cycle, choosing an hypothesis and carrying out experiments until it encounters a recursive call to itself or all the current hypotheses have been ruled out.

The system can also be viewed as carrying out a mixture of breadth-first and depth-first search through its hypothesis and experiment spaces. Under normal circumstances, KEKADA iterates

through each hypothesis and through each of the experiments generated to test these hypotheses;² one can view this behavior as a form of breadth-first search. However, when the result of an experiment is surprising, the system abandons this search and focuses attention on the newly found surprise. In essence, KEKADA forgets its previous goals and attends entirely to the new phenomenon, using it to generate a new set of hypotheses, and using these to guide further experimentation. One can view this control scheme as a form of depth-first search without backtracking, with the system being easily “distracted” by unexpected results. As we will see later, this counterintuitive strategy works quite well in the scientific domains we have examined.

3.3 KEKADA's Strategies

To constrain the search in the problem spaces, the system employs a small set of strategies to formulate a small number of good hypotheses and to suggest a few informative experiments. In addition to KEKADA's heuristic of focusing on surprises, these strategies are the system's basic source of power. We will now describe these strategies.

One strategy that KEKADA employs is to attempt to magnify a surprising phenomenon. For instance, if the system observes that switching on the electric current in a coil produces electric current in an adjacent coil, it would try changing the apparatus to increase the electric current in the second coil. If the surprising phenomenon has one or more apparatus variables associated with it, then the hypothesis generators suggest that the phenomenon may be magnified upon changing the value of the apparatus variables. When this hypothesis is chosen, the experiment proposers suggest a number of experiments in which one apparatus variable has a different value from that in the surprising phenomenon, and all the other variables have the same values.

If the experiment results in magnifying the surprising phenomenon,³ then all the future experiments characterizing the surprising phenomenon are carried out with the new set of values of the apparatus variable. This increases the chances of making crucial observations on further experimentation with the surprising phenomenon. For instance, suppose that, in one of its attempts to magnify the electric current, the system increases the length of the first coil to its maximum value, and this manipulation results in a significantly larger current in the second coil. KEKADA would then carry out all the further experiments with this longer coil, enabling it to make some crucial experimental observations.

A second strategy is a specific implementation of the divide and conquer method. If the surprising phenomenon is known to contain subprocesses, the hypothesis generators create the hypothesis that one of the subprocesses is behaving in an unexpected fashion. When the hypothesis is chosen, the

2. Kulkarni (1988) describes the conditions under which the system employs a different strategy.

3. In these experiments, the non-numeric variables can have all possible values, but the numeric variables can have only the maximum and the minimum values.

experiment proposers suggest a number of experiments on this subprocess in the same manner. If the results reveal that the subprocess is behaving in an unexpected manner, the experiment evaluators make a recursive call to KEKADA. For instance, suppose the system finds that alanine is not producing urea as expected, and it knows that this process has two subprocesses. In this case, the hypothesis generators would suggest the hypothesis that one of two subprocess is behaving in an unexpected manner and would invoke KEKADA to resolve the surprise.

A third strategy involves trying to assess the scope of the surprising phenomenon using domain-specific taxonomies. If one of the independent variables involved in the surprising phenomenon belongs to a general class, then the hypothesis generators would suggest a hypothesis that this phenomenon is exhibited generally by members of this class. When this hypothesis is chosen, the experiment proposers would select members of the class, based on cost and availability, and test experimentally whether the surprising phenomenon is also exhibited by them. When the number of members exhibiting this phenomenon exceeds the threshold value of three, the experiment evaluators interpret that the hypothesis is correct, generalize the description of the surprising phenomenon, and make a recursive call to KEKADA. In contrast, when the number of members that do not exhibit this surprising phenomenon exceeds three, the system inactivates this hypothesis. For example, when KEKADA studies the surprising phenomenon in which ornithine produces ammonia in kidney tissue, its hypothesis generators suggest a number of scope hypotheses. Upon carrying out various experiments, the system infers that the effect is exhibited generally by amino acids.

A fourth strategy involves finding the relevant factors in a surprising phenomenon. If the phenomenon has two or more independent entities, then the hypothesis generator would suggest a hypothesis that all the independent entities are not necessary to produce this effect. When this hypothesis is chosen, experiment proposers would suggest a number of experiments, each of which has all the independent entities in the phenomenon except for one. If one of these experiments exhibits the surprising effect, the experiment evaluators infer that the effect can be exhibited with fewer independent entities than those in the originally observed surprising phenomenon and then make a recursive call to KEKADA. For instance, suppose the system observes that, when electric current is switched on in an experimental setup, electric current appears in another coil. The system would then examine which factors are necessary to cause this effect. Now suppose it finds that switching electric current in a coil alone is sufficient to cause the effect. The system would then focus on this effect.

A fifth strategy involves looking for phenomena that are similar to the surprising phenomena in some way and then trying to find some relation between them. In particular, if two anomalous effects include the same variable, then the hypothesis generators suggest two hypotheses. The first hypothesis is that a larger class of values of this variable would exhibit anomalous effects. The experiment proposers and experiment evaluators associated with such a scope hypothesis were described earlier. The second hypothesis is that the two effects are part of a more complex

mechanism of the form $A \rightarrow B \rightarrow C$. The experiment proposer for this hypothesis is specific to the domain of metabolic biochemistry in the early 1900s. It suggests that one should measure the rates of formation of C from A and from B. If the second rate is slower than the first, then the experiment evaluators increment the implied-failure attribute of the confidence vector for the hypothesis. In contrast, if the first rate is slower than the second rate, then the experiment evaluators increment the implied-success attribute of the confidence vector for the hypothesis.

KEKADA employs these strategies, along with a few others, to create a small number of hypotheses. Each hypothesis has a priority level associated with it, and the system prefers hypotheses with higher priority levels. Furthermore, it chooses between hypotheses with the same priority level using the confidence vectors. Kulkarni (1988) describes both the strategies and the preference scheme in detail.

4. Rediscovery of Glutamine Synthesis

Now that we have described the representation, the control structure, and the processes in KEKADA, we will examine the behavior of the system on a particular problem from the history of science. In 1933, the biochemist Hans Krebs worked on the problem of understanding the nature of amino acid metabolisms. He established that the deamination of amino acids (i.e., the removal of the amino group from the amino acids) occurs by an oxidative reaction in the kidney, and not in the liver, as had been previously assumed. He also produced data on the deamination rates of various amino acids. Furthermore, he showed that glutamic acid combines with ammonia to produce glutamine, a substance that was not previously known to play any role in metabolism. The discovery of the glutamine reaction opened a whole set of new questions in metabolic biochemistry.

Here we examine KEKADA's behavior when it is given a problem similar to that faced by Hans Krebs. The problem can be stated as:

- *Given:* A surprising reaction in which ornithine produces ammonia in the presence of kidney tissue slices;
- *Given:* Three previously postulated reaction pathways about the deamination reaction (described later);
- *Given:* Background knowledge about the chemistry of various substances;
- *Do:* Revise the existing domain theory for amino acid metabolisms.

KEKADA's behavior on this research problem can be divided into three stages: characterization of the ornithine-in-kidney effect, study of the deamination reaction, and discovery of the glutamine reaction.

In the first stage, the system tries to characterize the surprising phenomenon in which ornithine produces ammonia in the kidney. Experiments reveal that other amino acids are also able to produce the effect, so KEKADA concludes that the phenomenon is a specific instance of the more general deamination reaction. During the second stage, the system carries out a variety of experiments that reveal details of the deamination reaction. In the process, the system comes across an unusual reaction in which the presence of arsenite increases the production of ammonia from the glutamic acid. In the final stage, KEKADA conjectures that glutamic acid combines with ammonia. The system verifies experimentally that this reaction occurs and that it produces glutamine as output. Below we describe this discovery process in greater detail.

4.1 Characterization of the Ornithine-in-Kidney Effect

In response to the surprise that ornithine produces ammonia in kidney, the hypothesis generators suggest a number of alternative explanations. One of the generators suggests assessing the scope of the phenomenon whenever the reactants in the observed phenomenon belong to a class of substances. In this case, ornithine belongs to the class of amino acids, amines, and carboxylic acids. Therefore, this generator suggests a hypothesis that the effect may be common to one of these classes. If the phenomenon has at least one apparatus variable associated with it, the *magnification generator* suggests attempting to magnify the effect. As ornithine has two amino groups and ammonia also has an amino group, another hypothesis generator suggests that ornithine is donating one of its amino groups to ammonia.

For reasons discussed in Kulkarni (1988), the preference scheme in KEKADA prefers the magnification hypothesis over other hypotheses. Thus the system attempts to magnify the observed effect. However, attempts to magnify the phenomenon by varying the apparatus variables fail, so next KEKADA decides to assess the scope of the surprising phenomenon. This leads to experiments on a number of amino acids, which reveal that other amino acids can also produce ammonia in the kidney. When four different amino acids are found to produce ammonia in the kidney, experiment evaluators generalize the original observation and make a recursive call to KEKADA.

4.2 Study of the Deamination Reaction

Now that the system has determined that the surprising phenomenon is exhibited by amino acids in general, the hypothesis generators suggest a number of new hypotheses. As the observed phenomenon shares a number of entities with deamination, one hypothesis generator suggests that the

observed phenomenon may be deamination and thus is one of the three previously hypothesized deamination reactions:

- Hydrolytic: $\text{R-CH-COOH} + \text{H}_2\text{O} = \text{R-CHOH-COOH} + \text{NH}_3$
- Reductive: $\text{R-CH(NH}_2\text{)-COOH} + \text{H}_2 = \text{R-CH}_2\text{-COOH} + \text{NH}_3$
- Oxidative: $2 \text{R-CH (NH}_2\text{)-COOH} + \text{O}_2 = 2 \text{R-CO-COOH} + 2 \text{NH}_3$

The original observation of the effect was specific to ornithine. In this case, KEKADA established the generality of the effect, which let it use general knowledge about amino acids to create additional hypotheses. The system first decides to verify the oxidative hypothesis, but it must choose an amino acid to carry out an experiment. KEKADA decides to carry out the experiment on alanine because this substance is both reactive and cheap. Here we see another advantage of the strategy of assessing the generality of a surprise. If the phenomenon turns out to be general, the experimenter has more choice in choosing a variable in the phenomenon. To verify the oxidative reaction, KEKADA carries out experiments on alanine and oxygen together. The results are consistent with the chemistry of the oxidative reaction.

The system next decides to gather more data on the deamination of amino acids. After carrying out experiments on a number of other amino acids, it carries out an experiment on glutamic acid. KEKADA expects the glutamic acid to deaminate in a similar way to other amino acids, as expectation setters in the system use the information of a class to set expectations about specific substances. However, the system finds that the rate of production of keto acid from glutamic acid is lower than expected. KEKADA notes this as a surprise and makes glutamic acid reaction the focus of attention.

As the system focuses on this surprise, the hypothesis generators associated with various strategies suggest a number of hypotheses. Using its preference scheme, KEKADA first chooses the magnification hypothesis. It then attempts unsuccessfully to magnify the phenomenon by changing the tissue and the aerobic conditions.

Next, the system selects the hypothesis that addition of an inhibitor substance would selectively block a side reaction consuming the keto acid. Expectation setters use the chemistry-specific knowledge about the action of inhibitors to make predictions about the results of the experiment. KEKADA does not predict whether the rate of production of keto acid will increase, but it does expect that the rate of production of ammonia will not be affected. When the reaction is carried out with arsenite, a particular inhibitor, the rate of production of both keto acid and ammonia increases. This violates the expectations, leading the system to focus its attention on this newly found surprising phenomenon.

4.3 Discovery of the Glutamine Reaction

At this point, the hypothesis generators associated with the various strategies in KEKADA suggest a number of different hypotheses. Among others, these include the following:

- The phenomenon may be common to some larger class, such as the class of carboxylic acids, amino acids, or amines.
- Ammonia may be reacting with a reactant in a side reaction.

Based on its preference scheme, KEKADA decides to test the hypothesis that the phenomenon may be common to the class of carboxylic acids. However, after getting negative results for aspartic acid and other carboxylic acids, the system reduces its confidence in this hypothesis. At this point, the system decides to consider the hypothesis that ammonia is reacting with one of the other reactants in a side reaction. When it carries out a reaction with glutamic acid and ammonia, glutamine is produced. Thus the system discovers the important glutamine reaction.

KEKADA's operation on the ornithine-in-kidney problem produced a number of interesting results. It established that deamination of amino acids occurs by an oxidative reaction in the kidney, and not in the liver, as had been previously assumed. It produced data on deamination rates of various amino acids. Furthermore, it showed that glutamic acid combines with ammonia to produce glutamine, a substance that was not previously known to play any role in metabolisms. The basic source of KEKADA's power was its ability to create a small number of good hypotheses and experiments and to focus on surprises.

4.4 Summary of the Discovery Process

Table 4 shows the state of KEKADA's knowledge at five steps, labeled S1, S2, S3, S4, and S5, in the operation just described. Each step is associated with the knowledge KEKADA has at that point in the run. Initially (S1), the system lacks the knowledge about amino acid metabolisms that it would acquire at the end of the run (S5). At S1, it does not know the exact nature of the deamination reaction, that it occurs in the kidney, or that it involves a glutamine reaction.

At this point, the system attends to the surprise that ornithine produces ammonia in the kidney. This leads it to carry out experiments mainly on various amino acids in the kidney tissue slices. In doing so, it searches a small part of the problem space that is rich in informative experiments. Thus, the surprise about the ornithine reaction in the kidney constrains KEKADA's search. The system then discovers that amino acids deaminate in the kidney and focuses on this general effect (S2); at this point, it has generalized the ornithine effect that it knew at S1.

At a later stage, when the system detects the surprising glutamic acid effect, it attends to this effect instead of to the previously found surprises (S3). This leads it to carry out a number of

Table 4. KEKADA's knowledge at various points in the discovery of glutamine synthesis.

S1:	The system focuses attention on the ornithine-in-kidney effect.
	ornithine \rightarrow ammonia in kidney
S2:	The system focuses attention on the amino-acids-in-kidney effect.
	amino-acid \rightarrow ammonia in kidney
S3:	The system focuses attention on the glutamic acid effect.
	amino-acid + oxygen \rightarrow keto-acid + ammonia
	glutamic acid \rightarrow alpha-keto-glutarate ($<< 1$) + ammonia (< 1)
S4:	The system focuses attention on the arsenite effect.
	amino-acid + oxygen \rightarrow keto-acid + ammonia
	glutamic acid \rightarrow alpha-keto-glutarate ($<< 1$) + ammonia (< 1)
	glutamic acid + arsenite \rightarrow alpha-keto-glutarate + ammonia
S5:	The system has discovered the glutamine reaction.
	amino-acid + oxygen \rightarrow keto-acid + ammonia
	glutamic acid \rightarrow alpha-keto-glutarate ($<< 1$) + ammonia (< 1)
	glutamic acid + arsenite \rightarrow alpha-keto-glutarate + ammonia
	glutamic acid + ammonia \rightarrow glutamine

experiments on glutamic acid. Again, the surprise about glutamic reaction constrains the system's search to a smaller problem space that contains the glutamine reaction. Finally, the system discovers this reaction (S5). In summary, KEKADA employs a greedy strategy of focusing on a surprise as soon as it is detected, and this constrains its search.

5. Discussion

In this section, we evaluate KEKADA's performance and compare the system with other programs. In addition, we identify the limitations of the system and the ways it could be improved.

5.1 Evaluation of KEKADA's Performance

A scientist is judged by the results of his or her research, and KEKADA can similarly be judged by the discoveries it has made. We have tested the system on a number of different problems. In the previous section, we discussed the general problem of amino acid metabolisms and how KEKADA

responds to this problem. We also tested the system on the problem of urea synthesis in the body, which had been an open problem until 1932. In that year, the biochemist Hans Krebs showed that urea is synthesized in a cyclic mechanism, which KEKADA was able to rediscover. Another historical problem involved elucidating the structure of common alcohol. In 1853, Williamson showed that this substance has an ethyl group attached to an hydroxyl group, and KEKADA was able to rediscover this structure. In yet another run, the system was able to rediscover magneto-electric induction.

These successes demonstrate the generality of the overall system, but it is also important that a system's components themselves be general. KEKADA has 43 heuristics, of which 28 are domain independent. The remaining 15 are specific to a domain, such as biochemistry, but none is specific to a particular problem, such as urea synthesis. The system used 31 of its 43 heuristics in solving more than one task in the runs.

The generality of the system is also implied by its psychological plausibility. KEKADA's behavior matches very well against the behavior of scientists; Kulkarni and Simon (1988) have shown that the system is a good model of the heuristics Hans Krebs used in his discovery. The generality of KEKADA's heuristics and their applicability to a variety of problems suggest that the system would be effective in solving a wide class of tasks.

5.2 Relation to Other Research

It is useful to compare KEKADA with related work on experimentation and theory revision in machine discovery. For instance, Friedland (1979) studied the problem of producing a plan for an experiment and developed MOLGEN, a program that uses skeletal plan refinement to this end. In contrast, KEKADA carries out a complex series of experiments, changing its goals significantly along the way, but its specification of an experiment is abstract. Thus, a specification, such as "carry out an experiment with ornithine and ammonia with certain concentrations in liver using the tissue slice method," leaves out many details about how this experiment is to be carried out. The MOLGEN system can refine an abstract description into a detailed plan for execution of the experiment.

Two outgrowths of the original MOLGEN project were the programs called GENSIM and HYPGENE (Karp, 1989, 1990). These systems simulate a number of discoveries in molecular genetics. Given an experiment and a theory, GENSIM predicts the results of the experiment. If there is a mismatch between these predictions and the observations, HYPGENE infers modifications in the theory and the initial conditions of the experiment. To do this, it reasons backward from the differences between the predictions and the observations. Another novel feature of these systems is a powerful representation that is used to reason qualitatively about complex functional relations between variables. Both systems require prior knowledge to resolve an anomaly, and neither can carry out an extensive experimentation program.

Another program that explains anomalous experiments is ADEPT (Rajamoney, 1990). This program uses qualitative processes to represent theories and produces an explanation of an anomalous observation in terms of these processes. It then carries out experiments to confirm it. Like HYP-GENE, the system has a focused goal of explaining a particular phenomenon. In contrast, KEKADA has a more general goal of revising its theories by carrying out informative experiments. It uses the surprising effect as an opportunity to acquire more knowledge. Even in situations where a surprise cannot be explained using existing knowledge, KEKADA can be effective in revising the theories. Furthermore, it can detect new surprises and focus on these.

Other researchers (Klahr & Dunbar, 1988; Shrager & Klahr, 1986) have studied experimentation and theory revision in human subjects' understanding of complex devices. Their SDDS model accounts for fine-grained data in the form of think-aloud protocols; in contrast, KEKADA is based on coarse-grained data on historical discoveries. However, there are also commonalities between SDDS (Klahr & Dunbar, 1988) and KEKADA, such as their use of a dual-space framework.

5.3 Limitations and Future Work

Although we found that KEKADA's behavior matched well against that of scientists, our comparison also revealed some limitations of the system. We found that scientists appeared to employ a number of additional heuristics, some that were not useful on these particular problems but that would be useful on many other problems. As a result, they are able to solve significantly more problems than the system. One direction for future work is to develop a scientist's assistant by incorporating additional discovery mechanisms, along with a large body of knowledge about a given domain.

This would require representations that support both qualitative and quantitative reasoning, along with an architecture that can support a number of different mechanisms. A possible extension would be to adopt ideas from Nordhausen and Langley's IDS (1990), a system that integrates taxonomy formation, discovery of qualitative laws, and discovery of numeric laws. Their approach should support KEKADA's scope and factor-analysis strategies with a few changes, and it also provides an extended representation for experiments. Future work should attempt to integrate insights from the KEKADA and IDS systems.

6. Conclusion

In this chapter, we described KEKADA, a system that is capable of carrying out a complex series of experiments on problems from the history of science. The system incorporates a set of experimentation strategies that were extracted from the traces of the scientists' behaviors. KEKADA focuses on surprises to constrain its search, and uses its strategies to generate hypotheses and to carry out experiments. Some strategies are domain independent, whereas others incorporate knowledge of

a specific domain. The domain-independent strategies include magnification, determining scope, divide and conquer, factor analysis, and relating different anomalous phenomena. KEKADA represents an experiment as a set of independent and dependent entities, with apparatus variables and a goal. It represents a theory either as a sequence of processes or as abstract hypotheses.

In this chapter, we described KEKADA's response to a particular problem in biochemistry. On this and other problems, the system is capable of carrying out a complex series of experiments to refine domain theories. Analysis of the system and its behavior on a number of different problems has established its generality, but it has also revealed the reasons why the system, in its present form, would not be a good experimental scientist. Nevertheless, we believe our work advances the state of research on scientific discovery by proposing a set of computational strategies that can be applied to a wide variety of domains.

Acknowledgements

Editorial comments by Pat Langley have greatly improved the presentation of this chapter. Our work has benefited from discussions with Allen Newell, David Klahr, Tom Mitchell, Raul Valdes-Perez, and Peter Jansen. Thanks are also due to Kate McKusick for her comments on a draft of this chapter. The research was supported by Contract F33615-87-C-1499 from the Defense Advanced Research Projects Agency.

References

- Brownston, L., Farrell, R., Kant, E., & Martin, N. (1985). *Programming expert systems in OPS5: An introduction to rule-based programming*. Reading, MA: Addison-Wesley.
- Conant, J. (Ed.) (1957). *Harvard case histories in experimental science*. Cambridge, MA: Harvard University Press.
- Friedland, P. (1979). Knowledge based experiment design in molecular genetics. *Proceedings of the Sixth International Joint Conference on Artificial Intelligence* (pp. 285-287). Tokyo, Japan: Morgan Kaufmann.
- Karp, P. D. (1989). *Hypothesis formation and qualitative reasoning in molecular biology* (Technical Report KSL-89-52). Stanford, CA: Stanford University, Department of Computer Science.
- Karp, P. D. (1990). Hypothesis formation as design. In J. Shrager & P. Langley (Eds.), *Computational models of scientific discovery and theory formation*. San Mateo, CA: Morgan Kaufmann.
- Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science*, 12, 1-48.
- Kulkarni, D., & Simon, H. A. (1988). The processes of scientific discovery: The strategy of experimentation. *Cognitive Science*, 12, 139-175.
- Kulkarni, D. (1988). *The processes of scientific research: The strategy of experimentation*. Doctoral dissertation, Department of Computer Science, Carnegie Mellon University, Pittsburgh, PA.
- Nordhausen, B., & Langley, P. (1990). An integrated approach to empirical discovery. In J. Shrager & P. Langley (Eds.), *Computational models of scientific discovery and theory formation*. San Mateo, CA: Morgan Kaufmann.
- Rajamoney, S. (1990). A computational approach to theory revision. In J. Shrager & P. Langley (Eds.), *Computational models of scientific discovery and theory formation*. San Mateo, CA: Morgan Kaufmann.
- Shrager, J., & Klahr, D. (1986). Instructionless learning about a complex device: The paradigm and observations. *International Journal of Man-Machine Studies*, 25, 153-189.
- Simon, H. A., & Lea, G. (1974). Problem solving and rule induction: A unified view. In L. Gregg (Ed.), *Knowledge and cognition*. Hillsdale, NJ: Lawrence Erlbaum.